



# Monocular Visual-Inertial SLAM for ISMAR SLAM Challenge

Jie PAN

Shaozu CAO, Jie PAN, Jieqi SHI, Shaojie SHEN

Source Code: http://github.com/HKUST-Aerial-Robotics/VINS-Mono





#### **Monocular Visual-Inertial SLAM**

• System diagram



Source Code: http://github.com/HKUST-Aerial-Robotics/VINS-Mono





#### How to Use IMU?

- IMU integration
  - IMU has higher rate than camera
  - Cannot estimate all IMU states
  - Need to integration IMU measurements



Source Code: http://github.com/HKUST-Aerial-Robotics/VINS-Mono





## The Bad of IMU Integration in the Global Frame

- IMU integration in world frame
  - Requires global rotation at the time of integration







#### IMU Pre-Integration on Manifold

- IMU integration in the body frame of first pose of interests
  - IMU Integration without initialization
  - Can use any discrete implementation for numerical integration
  - Intuitive: "position" and "velocity" changes in a "free-falling" frame



Source Code: http://github.com/HKUST-Aerial-Robotics/VINS-Mono





\_



#### IMU Pre-Integration on Manifold

- Uncertainty propagation on manifold
  - Derive the error state model for the IMU pre-integration dynamics

$$\begin{split} \begin{bmatrix} \delta \dot{\alpha}_{t}^{b_{k}} \\ \delta \dot{\beta}_{t}^{b_{k}} \\ \delta \dot{\beta}_{t}^{b_{k}} \\ \delta \dot{\theta}_{t}^{b_{k}} \\ \theta \dot$$

#### Covariance matrix for pre-integrated IMU measurements 🧖

Source Code: http://github.com/HKUST-Aerial-Robotics/VINS-Mono





#### IMU Pre-Integration on Manifold

- Pre-integrated IMU measurement model
  - Describes the spatial and uncertainty relations between two states in the local sliding window



Source Code: http://github.com/HKUST-Aerial-Robotics/VINS-Mono

@2019 HKUST Aerial Robotics Group | http://uav.ust.hk





#### Vision Front-End

- Simple feature processing pipeline
  - Harris corners...
  - KLT tracker...
  - Track between consecutive frames, flow back
  - RANSAC for preliminary outlier removal



Source Code: http://github.com/HKUST-Aerial-Robotics/VINS-Mono





#### **Monocular Visual-Inertial SLAM**

• System diagram



Global Pose Graph Optimization and Map Reuse

Source Code: http://github.com/HKUST-Aerial-Robotics/VINS-Mono





• Nonlinear graph optimization-based, tightly-coupled, sliding window, visual-inertial bundle adjustment



Source Code: http://github.com/HKUST-Aerial-Robotics/VINS-Mono





- Nonlinear graph-based optimization
  - Optimize position, velocity, rotation, IMU biases, inverse feature depth, and camera-IMU extrinsic calibration simultaneously:

$$\begin{aligned} \mathcal{X} &= \begin{bmatrix} \mathbf{x}_0, \, \mathbf{x}_1, \, \cdots \, \mathbf{x}_n, \, \mathbf{x}_c^b, \, \lambda_0, \, \lambda_1, \, \cdots \, \lambda_m \end{bmatrix} \\ \mathbf{x}_k &= \begin{bmatrix} \mathbf{p}_{b_k}^w, \, \mathbf{v}_{b_k}^w, \, \mathbf{q}_{b_k}^w, \, \mathbf{b}_a, \, \mathbf{b}_g \end{bmatrix}, k \in [0, n] \\ \mathbf{x}_c^b &= \begin{bmatrix} \mathbf{p}_c^b, \, \mathbf{q}_c^b \end{bmatrix}, \end{aligned}$$

Minimize residuals from all sensors







- IMU measurement residual
  - Additive for "position" and "velocity" changes, and biases
  - Multiplicative for incremental rotation

 $\mathbf{r}_{\mathcal{B}}(\hat{\mathbf{z}}_{b_{k+1}}^{b_{k}}, \mathcal{X}) = \begin{bmatrix} \delta \alpha_{b_{k+1}}^{b_{k}} \\ \delta \beta_{b_{k+1}}^{b_{k}} \\ \delta \beta_{b_{k+1}}^{b_{k}} \\ \delta b_{a} \\ \delta b_{g} \end{bmatrix} = \begin{bmatrix} \mathbf{R}_{w}^{b_{k}}(\mathbf{p}_{b_{k+1}}^{w} - \mathbf{p}_{b_{k}}^{w} + \frac{1}{2}\mathbf{g}^{w}\Delta t_{k}^{2} - \mathbf{v}_{b_{k}}^{w}\Delta t_{k}) - \hat{\alpha}_{b_{k+1}}^{b_{k}} \\ \mathbf{R}_{w}^{b_{k}}(\mathbf{v}_{b_{k+1}}^{w} + \mathbf{g}^{w}\Delta t_{k} - \mathbf{v}_{b_{k}}^{w}) - \hat{\beta}_{b_{k+1}}^{b_{k}} \\ 2 \left[ \mathbf{q}_{b_{k+1}}^{w^{-1}} \otimes \mathbf{q}_{b_{k}}^{w} \otimes \widehat{\mathbf{b}}_{b_{k+1}}^{b_{k+1}} \right]_{xyz} \\ \mathbf{b}_{ab_{k+1}} - \mathbf{b}_{ab_{k}} \\ \mathbf{b}_{wb_{k+1}} - \mathbf{b}_{wb_{k}} \end{bmatrix}$  IMU:  $\mathbf{X}_{0}$   $\mathbf{X}_{1}$   $\mathbf{X}_{1}$   $\mathbf{f}_{2}$ 

Source Code: http://github.com/HKUST-Aerial-Robotics/VINS-Mono

@2019 HKUST Aerial Robotics Group | http://uav.ust.hk

IMU pre-integration "blocks"





- Vision measurement residual
  - Pixel reprojection error
  - Inverse depth model, at least 2 observations per feature, first observation to define feature direction



Source Code: http://github.com/HKUST-Aerial-Robotics/VINS-Mono





- Marginalization
  - Bound computation complexity to a sliding window of states
  - Basic principles:
    - Add all frames into the sliding window, and remove non-keyframes after the nonlinear optimization
    - keep as many keyframes with sufficient parallax as possible
    - Maintain matrix sparsity by throwing away visual measurements from nonkeyframes



Source Code: http://github.com/HKUST-Aerial-Robotics/VINS-Mono





• Marginalization via Schur complement on information matrix



Source Code: http://github.com/HKUST-Aerial-Robotics/VINS-Mono

@2019 HKUST Aerial Robotics Group | http://uav.ust.hk





- Solving the nonlinear system
  - Minimize residuals from all sensors

$$\min_{\mathcal{X}} \left\{ \left\| \mathbf{r}_p - \mathbf{H}_p \mathcal{X} \right\|^2 + \sum_{k \in \mathcal{B}} \left\| \mathbf{r}_{\mathcal{B}}(\hat{\mathbf{z}}_{b_{k+1}}^{b_k}, \mathcal{X}) \right\|_{\mathbf{P}_{b_{k+1}}^{b_k}}^2 + \sum_{(l,j) \in \mathcal{C}} \left\| \mathbf{r}_{\mathcal{C}}(\hat{\mathbf{z}}_l^{c_j}, \mathcal{X}) \right\|_{\mathbf{P}_l^{c_j}}^2 \right\}$$

- Linearize (to Ax=b), solve, and iterate until time budget is reached
- Ceres Solver (<u>http://ceres-solver.org/</u>)





#### **Monocular Visual-Inertial SLAM**

• System diagram



------

Source Code: http://github.com/HKUST-Aerial-Robotics/VINS-Mono





- Speeding up
  - The sliding window monocular visual-inertial bundle adjustment runs at 10Hz
  - Motion-only visual-inertial bundle adjustment to boost up the state estimation 30Hz
  - IMU forward propagation to boost to 400Hz



Source Code: http://github.com/HKUST-Aerial-Robotics/VINS-Mono





- Motion-only visual-inertial bundle adjustment
  - Optimize position, velocity, rotation in a smaller windows, assuming all other quantities are fixed

 $\begin{aligned} \mathcal{X} &= \begin{bmatrix} \mathbf{x}_0, \ \mathbf{x}_1, \dots \cdot \mathbf{x}_n, \ \mathbf{x}_c^b, \lambda_0, \lambda_1, \dots \lambda_m \end{bmatrix} \\ \mathbf{x}_k &= \begin{bmatrix} \mathbf{p}_{b_k}^w, \ \mathbf{v}_{b_k}^w, \ \mathbf{q}_{b_k}^w, \ \mathbf{b}_a, \ \mathbf{b}_g \end{bmatrix}, k \in [0, n] \\ \mathbf{x}_c^b &= \begin{bmatrix} \mathbf{p}_c^b, \ \mathbf{q}_c^b \end{bmatrix}, \end{aligned}$ 



Prior in cost function is ignored



- Also solved using the Ceres Solver





#### **Monocular Visual-Inertial SLAM**

• System diagram



Global Pose Graph Optimization and Map Reuse

Source Code: http://github.com/HKUST-Aerial-Robotics/VINS-Mono





- Very, very, very important for monocular visual-inertial systems
- Assumption 1: known camera-IMU extrinsic calibration during initialization
  - Does not need to be very accurate
  - Extrinsic calibration is refined in later nonlinear optimization
- Assumption 2: known accelerometer and gyroscope biases during initialization
  - Use zero values at power-up
  - Use prior values during failure recovery
  - Reasonable assumption due to slow varying nature of biases
- Pipeline
  - Monocular vision-only SFM in a local window
  - Visual-inertial alignment





- Monocular vision-only structure-from-motion (SfM)
  - In a small window (10 frames, 1sec)
  - Up-to-scale, locally drift-free position estimates
  - Locally drift-free orientation estimates
  - Not aligned with gravity

IMU is not used in this step





- Visual-inertial alignment
  - Estimates velocity of each frame, gravity vector, and scale
    - Note the coordinate frames



Source Code: http://github.com/HKUST-Aerial-Robotics/VINS-Mono





• Visual-inertial alignment



- Solve a linear system
  - Scale and rotate the vSfM

$$\begin{aligned} \mathcal{X}_{I} &= \left[\mathbf{v}_{b_{0}}^{c_{0}}, \, \mathbf{v}_{b_{1}}^{c_{0}}, \, \cdots \, \mathbf{v}_{b_{n}}^{c_{0}}, \, \mathbf{g}^{c_{0}}, \, s\right] \\ \min_{\mathcal{X}_{I}} \sum_{k \in \mathcal{B}} \left\| \hat{\mathbf{z}}_{b_{k+1}}^{b_{k}} - \mathbf{H}_{b_{k+1}}^{b_{k}} \mathcal{X}_{I} \right\|^{2} \end{aligned}$$



Source Code: http://github.com/HKUST-Aerial-Robotics/VINS-Mono





#### **Monocular Visual-Inertial SLAM**

• System diagram



Global Pose Graph Optimization and Map Reuse

Source Code: http://github.com/HKUST-Aerial-Robotics/VINS-Mono





#### Loop Closure

- Loop detection
  - Describe features by BRIEF
    - Extract new FAST features
      - (500, only use for loop detection)
  - Query Bag-of-Word (DBoW2)
    - Return loop candidates





Calonder, Michael, et al. "Brief: Binary robust independent elementary features." *Computer Vision–ECCV 2010* (2010): 778-792. Gálvez-López, Dorian, and Juan D. Tardos. "Bags of binary words for fast place recognition in image sequences." *IEEE Transactions on Robotics* 28.5 (2012): 1188-1197.

Source Code: http://github.com/HKUST-Aerial-Robotics/VINS-Mono





#### Loop Closure

- Feature Retrieving
  - Try to retrieve matches for features that are used in the VIO
  - BRIEF descriptor match
  - Geometric check
    - Fundamental matrix test with RANSAC
    - At least 30 inliers
- Output:
  - Loop closure frames with known pose
  - Feature matches between VIO frames and loop closure frames



Source Code: http://github.com/HKUST-Aerial-Robotics/VINS-Mono





#### **Monocular Visual-Inertial SLAM**

• System diagram



Global Pose Graph Optimization and Map Reuse

Source Code: http://github.com/HKUST-Aerial-Robotics/VINS-Mono





# Monocular Visual-Inertial Odometry with Relozalization



Source Code: http://github.com/HKUST-Aerial-Robotics/VINS-Mono



電子及計算機工程學系 DEPARTMENT OF ELECTRONIC & COMPUTER ENGINEERING



### Monocular Visual-Inertial Odometry with Relozalization

- Relocalization
  - Visual measurements for tightly-coupled relocalization
    - Observation of retrieved features in loop closure frames
    - Poses of loop closure frames are constant
    - No increase in state vector dimension for relocalization
    - Allows multi-constraint relocalization



Source Code: http://github.com/HKUST-Aerial-Robotics/VINS-Mono







#### **Monocular Visual-Inertial SLAM**

• System diagram



Global Pose Graph Optimization and Map Reuse

Source Code: http://github.com/HKUST-Aerial-Robotics/VINS-Mono





| Seq Name | APE(mm) | RPE(mm) | ARE(deg) | RRE(deg) | Badness | InitQuality |
|----------|---------|---------|----------|----------|---------|-------------|
| C0       | 67.812  | 25.907  | 3.529    | 0.635    | 7.697   | 3.475       |
| C1       | 47.356  | 1.889   | 1.682    | 0.141    | 4.691   | 6.871       |
| C2       | 69.143  | 14.785  | 2.049    | 0.321    | 5.879   | 3.711       |
| C3       | 27.834  | 5.182   | 1.549    | 0.403    | 6.261   | 1.514       |
| C4       | 66.927  | 21.137  | 0.878    | 0.123    | 1.494   | 7.033       |
| C5       | 17.568  | 3.926   | 1.619    | 0.156    | 10.617  | 1.397       |
| C6       | 49.625  | 10.866  | 1.743    | 0.284    | 1.922   | 5.526       |

Source Code: http://github.com/HKUST-Aerial-Robotics/VINS-Mono





| Seq Name | APE(mm) | RPE(mm) | ARE(deg) | RRE(deg) | Badness | InitQuality |
|----------|---------|---------|----------|----------|---------|-------------|
| C7       | 14.735  | 2.676   | 0.805    | 0.130    | 9.699   | 1.848       |
| C8       | 47.193  | 6.187   | 3.530    | 1.040    | 4.058   | 1.341       |
| С9       | 20.167  | 3.180   | 1.619    | 0.318    | 15.422  | 2.264       |
| C10      | 40.418  | 11.506  | 4.993    | 0.212    | 10.038  | 3.225       |
| C11      | 31.688  | 6.602   | 1.726    | 0.589    | 9.659   | 1.498       |
| D8       | 29.524  | 8.242   | 2.936    | 0.818    | 21.139  | 1.257       |
| D9       | 14.064  | 1.484   | 1.896    | 0.180    | 0.218   | 0.910       |
| D10      | 152.871 | 26.105  | 12.285   | 8.072    | 32.598  | 0.543       |

Source Code: http://github.com/HKUST-Aerial-Robotics/VINS-Mono







Source Code: http://github.com/HKUST-Aerial-Robotics/VINS-Mono











Source Code: http://github.com/HKUST-Aerial-Robotics/VINS-Mono





@2019 HKUST Aerial Robotics Group | http://uav.ust.hk











Source Code: http://github.com/HKUST-Aerial-Robotics/VINS-Mono



電子及計算機工程學系 DEPARTMENT OF ELECTRONIC & COMPUTER ENGINEERING



# Thanks!

Questions?

Source Code: http://github.com/HKUST-Aerial-Robotics/VINS-Mono